

# データサイエンス実習

[実習] 第1学年 前期 必修 2単位

《履修上の留意事項》 毎回、各自のPCを必携のこと。授業では、毎回、各自のPCを使って演習をしますので、忘れずにPCを持参してください。

《担当者名》 教授／二瓶 裕之  
教授／塚本容子、講師／池森 康裕、講師／西牧 可織、助教／米田 龍大

## 【概要】

「数理・アルゴリズム演習」で修得したスキルを活かし、データサイエンティストにとって、より実践的なデータ処理やアルゴリズムの実装に取り組みます。実習では、臨床データや医療統計データを扱いながら、データベース設計・管理のスキルを応用します。また、Pythonを用いたデータ操作、基本的なデータ分析手法を実装して、医療データサイエンスにおける課題解決能力と瞬発力を養います。

第一回授業は、5名すべての教員が担当をします。塚本容子はデータサイエンスと看護のトピックス、池森康裕はデータサイエンスと福祉・介護（VRや介護ロボットなど）のトピックスについて講義して、医療のフィールドにおけるデータサイエンスの活用状況の概論を講義します。これを踏まえて、二瓶裕之、西牧可織、米田龍大の3名が、医療のフィールドにおいて活用されるデータサイエンスの具体的な技術について講義します。第二回授業以降は、二瓶裕之、西牧可織、米田龍大の3名で担当します。各回の授業開始冒頭において、二瓶裕之が実習課題の解説を担当します。解説の後、実習課題に取り組むときには、西牧可織と米田龍大が、教室内を巡回しながら、学生の質問等に対応します。

(オムニバス方式／全30回) (共同)

(二瓶裕之・塚本容子・西牧可織・池森康裕・米田龍大／1回) (共同)

(二瓶裕之・西牧可織・米田龍大／29回) (共同)

## 【学修目標】

- ☆臨床データや医療統計データから基本統計量を算出し、結果を説明できる
- ☆データ型や変数設計を分類し、適切な型選択や設計ができる
- ☆Pythonやデータベースを用いたデータの整理・検索・集計・可視化を操作できる
- ☆現場課題に即したアルゴリズムやデータ処理関数を設計・実装し、再利用の意義を説明できる
- ☆実践的なデータ分析を通じて、課題発見・提案・振り返りを表現できる

## 【学修内容】

回	テーマ	授業内容および学修課題	担当者
1	データサイエンス活用の概論	看護・福祉・介護分野でのデータサイエンス活用状況の概論	担当者全員
2	臨床データの基本統計量算出	看護・福祉・介護分野でのデータサイエンス活用状況の概論、医療データから人工知能を構築する手順を体験してみる。	二瓶裕之・西牧可織・米田龍大
3～4	データの変数設計・型選択	変数型・データ型を分類し、最適な型設計ができる。	二瓶裕之・西牧可織・米田龍大
5～6	演算処理の実装	各種のアルゴリズムに基づく演算処理を実装できる。	二瓶裕之・西牧可織・米田龍大
7～8	条件抽出によるデータ解析	データから特定条件のリスト等を抽出するなどのアルゴリズム設計ができる。	二瓶裕之・西牧可織・米田龍大
9～10	データ集計・異常検出	データセットに対し、集計・統計値・異常値検出を実装し、その有用性を評価できる。	二瓶裕之・西牧可織・米田龍大
11～12	現場課題に応じた関数化・再利用	データ処理関数を設計・実装し、再利用方法と意義を説明できる。	二瓶裕之・西牧可織・米田龍大
13～14	配列コレクション	大規模なデータセットに対して、ソートや検索などの操作ができる。	二瓶裕之・西牧可織・米田龍大
15～16	文字列処理	文字列データをクリーニング・コーディングし、正規化処理を実施できる。	二瓶裕之・西牧可織・米田龍大

17～18	関数化・再利用	データ処理関数を設計・実装し、再利用方法と意義を説明できる。	二瓶裕之・西牧可織・米田龍大
19～20	データの入出力・保存	ファイル保存・読み出しを Python で実施できる。	二瓶裕之・西牧可織・米田龍大
21～22	デバッグによるエラー解決	エラーを発見・修正し、再発防止策を考察できる。	二瓶裕之・西牧可織・米田龍大
23～24	ライブラリ・データベース連携の実践	データの整理・検索・集計にデータベースやライブラリを組み合わせる操作・管理できる。	二瓶裕之・西牧可織・米田龍大
25～26	データの集計・ソート・検索	必要な情報を抽出・検索・集計するアルゴリズムを設計・評価できる。	二瓶裕之・西牧可織・米田龍大
27～28	統計データの可視化と解釈	統計データを可視化できる。	二瓶裕之・西牧可織・米田龍大
29～30	総合課題	データ分析の全工程を振り返り、課題への提案・振り返りができる。	二瓶裕之・西牧可織・米田龍大

#### 【授業実施形態】

面接授業

※授業実施形態は、各学部（研究科）、学環、学校の授業実施方針による

#### 【評価方法】

授業参加態度（PBL の回には討議の参加態度）30%と毎回(15回)の授業で作成する課題の評価を70%で評価する。

#### 【教科書】

「Python 実践データ分析 100本ノック 第3版」下山 輝昌（著）、松田 雄馬（著）、三木 孝行（著） 秀和システム

#### 【備考】

- ・ 本学 DX 推進サイト (<https://dx.hoku-iryo-u.ac.jp/>) に公開している資料を使用する。
- ・ 東京大学 数理・情報教育研究センター ([http://www.mi.u-tokyo.ac.jp/6university\\_consortium.html](http://www.mi.u-tokyo.ac.jp/6university_consortium.html)) に公開されているリテラシーレベル教材と応用基礎レベル教材も使用する。

#### 【学修の準備】

- ・ 予習としては、自分で所有しているパソコンを利用して事前学修の項目を確認すること(45分)。
- ・ 復習としては、授業で作成した課題を再度作成するなどの事後学習を行うこと(45分)。
- ・ SGD や PBL では、事前に具体的な調査資料を配布するので、個々で調べて授業に臨むこと(45分)。授業終了後は、グループ討議の結果を自分なりに振り返り、授業中に作成したプロダクトを再度作成するなどの事後学修（復習）を行うこと(45分)。

#### 【ディプロマ・ポリシー(学位授与方針)との関連】

1. 一般的なデータサイエンティストとしてのデータ分析・解析・AI などにかかるスキルにより瞬発力をもって課題を解決でき、加速度的に広がる生成 AI などの先進的技術を応用できる実践能力を身につけている。

#### 【アクティブ・ラーニング】

導入している

#### 【その他】

この科目は主要授業科目に設定している